

# Weekly Report

Yuxin Ma

12.02.2013 - 12.08.2013

## Intro

This week Visual SVM project is going on. I also spend several days on reviewing undergraduate InfoVis course.

## Research

**Viewport selection** One of the most important problem in tour projection is view selection. Views are determined by the transformation matrix set by user him/herself or by automated approaches. However, unlike general projection, the most significant difference is that the view port should be parallel to the separating plane. Following is the mathematical representation of this issue:

Let  $\mathcal{X} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m\}$  be the orthonormal basis of the original data space, and  $\mathbf{n}_1$  be the normal vector of the SVM separating plane, then  $\mathbf{n}_1$  is one of the vectors in the orthonormal basis of the projection plane. Denote another vector in the orthonormal basis as  $\hat{\mathbf{n}}$ , and the projection plane can be represented as  $\text{span}(\mathbf{n}_1, \hat{\mathbf{n}})$ . Also, a space can be generated by  $\mathbf{n}_1$ , denoted as  $\text{span}(\mathbf{n}_1, \mathbf{n}_2, \dots, \mathbf{n}_m) = \mathcal{X}$ . Then  $\hat{\mathbf{n}}$  can be rewrited as an linear combination of  $\mathbf{n}_2, \dots, \mathbf{n}_m$ , namely

$$\hat{\mathbf{n}} = \sum_{i=2}^m w_i \mathbf{n}_i$$

For  $\hat{\mathbf{n}}$  is normal, we have

$$\begin{aligned} |\hat{\mathbf{n}}| &= \left| \sum_{i=2}^m w_i \mathbf{n}_i \right| = 1 \\ \Rightarrow \sum_{i=2}^m w_i^2 &= 1 \end{aligned}$$

For Radviz method, it requires that the summation of weights is 1 but not the square-summation. One possible solution is to transform each  $w_i^2$  into  $\hat{w}_i$ , and use the new weights in Radviz. This solution needs further investigation.

There are two styles of dimension weight adjustment (i.e. selecting transformation matrix):

- Manual adjustment: For each dimension<sup>1</sup>, a “spin wheel” is provided for controlling “rotation” on the dimension (just like controlling the volumn button on a radio). There

---

<sup>1</sup>The word “dimension” is referred to as original dimensions of the data.

should be some papers about manually adjusting projections for our method. I will do a survey on this.

- Automated pre-computed projections: Similar to projection pursuit, several candidate views are generated first according to some metrics. Then user is involved in the exploration of candidate views to find informative ones.

The following challenge is to find which kind of projections is valuable for users to gain more information on data distribution and support vector distribution/importance and dimension importance. The metric might be an optimization equation.

In addition, there should be a relation between feature selection and view selection. For data explanation, different feature selection methods represents different bias, which can be paired with corresponding view selection methods.

**Implementation** This week I paid more attention on the previous codes of the server side. Now the functions of Sessions, Model Configurations, Datasets and Models are ready to use for the browser side. Next week I plan to spend two days for rewriting the interface framework and migrating previous projection and control views to the new one.

## Next Plans

- Push the process of implementation.